

The embodiments of the invention in which an exclusive property or privilege is claimed are defined as follows:

1. A computer-implemented method for determining the legibility of text in an image of a page, the method comprising:

- (a) obtaining an image of a page having text therein;
- (b) performing text recognition on text in the page image and producing a measure of the text;
- (c) comparing the measure of the text in the page image with a measure derived from the page image itself; and
- (d) determining the legibility of the text in the page image based on the comparison of the text and page image measures.

2. The method of Claim 1, wherein performing text recognition comprises identifying a spatial dimension of the text in the page image.

3. The method of Claim 2, wherein the measure of the text in the page image is a measure of text height in the identified spatial dimension, and the measure derived from the page image is a measure of page image height in the same spatial dimension.

4. The method of Claim 3, wherein comparing the measure of text height to page image height includes dividing the text height by the page image height to produce a text-to-page height ratio.

5. The method of Claim 4, wherein determining the legibility of the text includes comparing the text-to-page height ratio to a threshold.

6. The method of Claim 5, wherein the text is determined to be not legible if the text-to-page height ratio does not satisfy the threshold.

7. The method of Claim 5, wherein text recognition is performed on text in multiple lines in the page image and each of the multiple lines of text produces a text-to-page height ratio, wherein each of the multiple text-to-page height ratios are compared to the threshold, and the legibility of the text in the page image is determined based on a percentage of the text-to-page height ratios that satisfy the threshold.

8. The method of Claim 5, wherein the threshold is a first threshold, the method further comprising comparing the text-to-page height ratio to a second threshold if the text-to-page height ratio satisfies the first threshold.

9. The method of Claim 8, wherein the text is determined to be legible if the text-to-page height ratio satisfies the second threshold.

10. The method of Claim 8, wherein the text is determined to be possibly legible if the text-to-page height ratio satisfies the first threshold but not the second threshold.

11. The method of Claim 8, wherein text recognition is performed on text in multiple lines in the page image and each of the multiple lines of text produces a text-to-page height ratio, wherein if a determined percentage of the multiple text-to-page height ratios satisfy the first threshold, the text-to-page height ratios are compared to the second threshold, and if a determined percentage of the text-to-page height ratios satisfy the second threshold, the text is determined to be legible.

12. The method of Claim 1, further comprising storing the page image for display if the text in the page image is determined to be legible.

13. The method of Claim 1, wherein if the text in the page image is determined to be not legible, a higher resolution image of the page is obtained.

14. The method of Claim 1, wherein performing text recognition includes identifying and counting words that comprise the text in the page image, and wherein producing a measure of the text includes calculating a measure of word density in the page image based on the number of words in the page image and the spatial size of the page image.

15. The method of Claim 14, wherein the measure derived from the page image comprises:

- (a) compressing the page image to form a compressed image file; and
- (b) determining the file size of the compressed image file.

16. The method of Claim 15, wherein determining the legibility of the text includes comparing the measure of word density to the file size of the compressed image file.

17. A computer-implemented method for displaying a page image based on a determined legibility of text in the page image, the method comprising:

- (a) obtaining an image of a page at a base resolution;
- (b) determining the legibility of text in the page image; and
- (c) if the text in the page image is determined to be legible, then displaying the page image, otherwise obtaining and displaying a new image of the page obtained at a resolution higher than the base resolution.

18. The method of Claim 17, wherein the base resolution page image is obtained from an original high resolution image of the page, and if the text in the base resolution page image is determined to be not legible, then the new page image is obtained from the original high resolution page image at a resolution higher than the base resolution.

19. The method of Claim 17, wherein determining the legibility of text in the page image includes:

- (a) performing text recognition on text in the page image and producing a measure of the text;
- (b) comparing the measure of the text in the page image with a measure derived from the page image itself; and
- (c) determining the legibility of the text in the page image based on the comparison of the text and page image measures.

20. The method of Claim 19, wherein performing text recognition comprises identifying a spatial dimension of the text in the page image, wherein the measure of the text in the page image is a measure of text height in the identified spatial dimension, and wherein the measure derived from the page image is a measure of page image height in the same identified spatial dimension.

21. The method of Claim 20, wherein comparing the measure of text height to page image height includes dividing the text height by the page image height to produce a text-to-page height ratio, and wherein determining the legibility of the text includes comparing the text-to-page height ratio to a threshold.

22. The method of Claim 19, wherein text recognition is performed on multiple lines of text in the page image.

23. The method of Claim 19, wherein performing text recognition includes identifying and counting words that comprise the text in the page image, wherein producing the measure of the text includes calculating a measure of word density in the page image based on the number of words in the page image and the spatial size of the page image, wherein the measure derived from the page image includes compressing the page image to form a compressed image file and determining the file size of the compressed image file, and wherein determining the legibility of the text includes comparing the measure of word density to the file size of the compressed image file.

24. A multistage method for automated determination of legibility of text in an image of a page, comprising:

- (a) obtaining an image of a page having text therein;
- (b) determining the legibility of the text by applying a first test of legibility to the text in the page image;
- (c) determining the legibility of the text by applying a second test of legibility to the text in the page image; and
- (d) storing the page image for display if the text in the page image is determined to be legible.

25. The method of Claim 24, wherein the page image is stored for display if the text in the page image is determined to be legible by either the first test or the second test.

26. The method of Claim 24, wherein the page image is stored for display if the text in the page image is determined to be legible based on the results of both the first test and the second test.

27. The method of Claim 24, wherein if the text in the page image is determined not legible based on the first test, the second test is not applied to the page image.

28. The method of Claim 24, wherein the result of the first test is one of legible, not legible, or possibly legible, and wherein the second test is applied if the result of the first test is possibly legible.

29. A computer-implemented system configured to determine the legibility of text in an image of a page, comprising a processor programmed to obtain an image of a page having text therein, perform text recognition on the text in the page image and produce therefrom a measure of the text, compare the measure of the text in the page image with a measure derived from the page image itself, and determine the legibility of the text in the page image based on the comparison of the text and page image measures.

30. The system of Claim 29, wherein the processor is programmed to measure the text in the page image by measuring a height of a line of text, and further measure the height of the page image to derive the measure of the page image.

31. The system of Claim 30, wherein the processor is programmed to determine the legibility of the text by comparing the measure of text height with the page image height.

32. The system of Claim 29, further comprising a memory in communication with the processor for storing the page image for display if the text in the page image is determined to be legible.

33. The system of Claim 29, wherein if the text in the page image is determined to be not legible, the processor is programmed to obtain a higher resolution image of the page.

34. The system of Claim 29, wherein the processor is programmed to identify and count words that comprise the text in the page image and produce a measure of the text by calculating a word density based on the number of words in the page image and a spatial size of the page image.

35. The system of Claim 34, wherein the processor is programmed to determine the legibility of the text by compressing the page image to form a compressed image file, determine the file size of the compressed image file, and compare the calculated word density with the file size of the compressed image file.

36. The system of Claim 29, further comprising a display in communication with the processor for displaying the page image based on a determined legibility of the text in the page image.

37. The system of Claim 36, wherein if the text in the page image is determined to be not legible, the processor is programmed to obtain and communicate to the display a higher resolution image of the page.

38. The system of Claim 29, further comprising a memory in communication with the processor, wherein the processor is programmed to determine the legibility of the text by applying a first test of legibility and a second test of legibility to the text in the page image, and store the page image in the memory for display if the text in the page image is determined to be legible.

39. The system of Claim 38, wherein the processor is programmed to store the page image for display if the text in the page image is determined to be legible by either the first or the second test of legibility.

40. The system of Claim 38, wherein the processor is programmed to store the page image for display if the text in the page image is determined to be legible based on the results of both the first and second tests of legibility.

41. The system of Claim 38, wherein if the processor determines the text in the page image is not legible based on the first test of legibility, the processor is programmed to not apply the second test of legibility to the page image.